1. Full Title:
Two papers are planned, one more applied and submitted to a "medical" journal; the other more methodological and submitted to an "epidemiological" journal. The applied article will focus on the impact of the bias (discussed below) on a single risk factor, most likely smoking, and discuss how the estimates of the impact are likely to be biased to produced underestimates of effects at older ages. The other paper will address the more general problem of the implication of changes in the correlational structure at different ages.

Applied Paper: **Is the estimated impact of cigarette smoking underestimated?**


Abbreviated Title (length 26):       Applied paper: Bias in smoking effect
                                         Methodological: Risk factor correlation

2. Writing Group (list individual with lead responsibility first):
Lead: George Howard

Department of Public Health Sciences
Bowman Gray School of Medicine
Winston Salem, NC 27157
Phone: (910) 716-2865     Fax: (910) 716-5425
Email: HOWARD@PHS.BGSM.WFU.EDU

Tim Morgan         Lynne Wagenknecht (CARDIA)         Greg Burke (CHS)            Lloyd Chambless               Aaron Folsom

3. Timeline:
Immediate

4. Rationale:
It has been suggested that the impact of traditional risk factors (including hypertension[1] and lipids[2]) on CVD outcomes may be smaller in elderly populations. A potential explanation of this finding is a change in the correlational structure of risk factors across age. Because of many reasons including the insulin resistance syndrome or life-style choices, risk factors tend to "cluster" or be correlated[3]. As such, at a young age a positive correlation may be seen. This paper will address the hypothesis that with increasing age a decreased correlation (or a negative correlation) between risk factors may be seen. For example, consider smoking, hypertension, diabetes, and hyperlipidemia
as risk factors underlying heart disease. At a young age, there can be a number of individuals in a population with two or more of these risk factors, and because of the correlation between these risk factors individuals with one risk factor (say smoking) will also tend to have other risk factors (hypertension, hyperlipidemia, and diabetes). However, at old age most individuals with multiple risk factors have been selectively removed from the population. So at an older age, individuals with one risk factor (say smoking) will tend not to have other risk factors (because all of the participants who smoke, and are hypertensive, hyperlipidemic, diabetics have died). Mathematically, in terms of the correlational structure between risk factors, this will be reflected in a decrease in the correlation between risk factors, or even a negative correlation between risk factors at older ages.

If this decrease in the correlational structure between risk factors proves to be the case, it will have a substantial impact on the estimated magnitude of impact of risk factors. At a young age, because of the positive correlation between risk factors the estimated of the impact of a risk factor will tend to be overstated. For example, those smoking also tend to be hypertensive, diabetic, and hyperlipidemic; and when the impact of smoking is estimated part of the impact of these other factors is partially included because of the positive correlational structure. However, at an older age because of the negative correlation structure the estimated impact of a risk factor will tend to bias estimates differentially across ages. For example, those elderly smokers, tend to be "protected" by not being hypertensive, diabetic and hyperlipidemic; hence, the estimated impact of smoking is reduced because of protection from the absence of other risk factors. This change in correlational structure will lead to the findings observed for a decreased impact of hypertension and hyperlipidemia among the elderly. It may be argued that the change in the correlational structure can be accounted for by covariate adjustment in multivariable analyses. However, this assumption implies that all covariates (known and unknown) are appropriately modeled, clearly a difficult assumption.

Combining the CARDIA, ARIC, and CHS cohorts, data will be available on the major cardiovascular risk factors over the ages from approximately 30 years to well over 80 years. The focus of the paper would be on the correlational structure of those risk factors that are part of the Framingham Risk Scale (FRS) [4], specifically smoking, hyperlipidemia (TC/HDL ratio), diabetes, and hypertension (systolic blood pressure). For each of these risk factors, a residual risk factor score will be calculated using the coefficients in the FRS to summarize the risk factor levels of the other three risk factors. The correlation between the risk factor and the residual risk score will then be calculated for participants in a 10-year wide sliding window. This 10-year window will then be moved through the age spectrum from the young ages to the older ages, and the change in the correlation between the risk factor and the residual risk score tabulated. This process will then be repeated for each of the risk factors in the FRS.

5. Main Hypothesis:
The correlational structure between risk factors will weaken or become negative at older ages, implying changes in the estimated association between risk factors and outcome measures.
6. Data (variables, time window, source, inclusions/exclusions):
CARDIA year 10 data set (ages 28 to 40), ARIC baseline data set (ages 45 to 65), and
CHS baseline data set (ages 65+). Within each, the variables used include basic risk
factor measures: diabetes status, systolic blood pressure, total cholesterol, HDLc,
cigarette smoking status, and age.

REFERENCES:
1. Langer RD, Ganiats TG, Barrett-Connor E. Paradoxical survival of elderly men with

2. Manolio TA, Pearson TA, Wenger NK, Barrett-Connor E, Payne GH, Harlan WR.
Cholesterol and heart disease in older persons and women. *Ann Epidemiol* 1992;2:161-
176.

1607.